

NEB-241-PUS.ST25.txt
 SEQUENCE LISTING

<110> New England Biolabs, Inc.
 Morgan, Richard
 Wilson, Geoffrey
 Lunnen, Keith
 Heiter, Daniel
 Benner, Jack
 Nfenfou, Celine
 Picone, Stephen

<120> A Novel Modular Type II Restriction Endonuclease, CspCI, and the
 Use of Modular Endonucleases for Generating Endonucleases with
 New Specificities

<130> NEB-241-PUS

<150> 60/555,796
 <151> 2004-03-24

<150> PCT/US05/09824
 <151> 2005-03-23

<160> 49

<170> PatentIn version 3.2

<210> 1
 <211> 61
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n=a,c, g or t

<400> 1
 ccccgaaaag tncacctga cgtgcaacct aggtggcacg tctaagaaac cattattatc 60
 a 61

<210> 2
 <211> 61
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n=a,c, g or t

<400> 2
 tgataataat ggtntcttag acgtgccacc taggttgac gtcaggtggc acttttcggg 60

g

61

<210> 3
 <211> 61
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> n=a,c,t or g

<400> 3
 ccccgaaaag tncacactga cgtgccaccc ggggtgcacg tctaagaaac cattattatc 60

a 61

<210> 4
 <211> 61
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n-a,c,g or t

<400> 4
 tgataataat ggtntcttag acgtgcaacc cgggtggcac gtcaggtggc acttttcggg 60

g 61

<210> 5
 <211> 4616
 <212> DNA
 <213> unknown

<220>
 <223> Citrobacter species 2144

<400> 5
 agatctgccca atactgtttc gacagcgcca cttaatctct tcaatttcgc gcaggttaga 60

tggcacttgt tcggaagagg cgtctgtaac tcggtctcaa gctgcgggat tgccccggta 120

ttttgctcat ccccttttaa ttcaacgatc atctgctgaa aagtcagacc ctcacgatag 180

tcagaggcca gcttctgttc ctcagatgcc agccctttca gatcgggtctt acccgtctct 240

agttgttggg ttgccccatc aagagtgcgg cgtttttctg ccaattgagt agctttttaca 300

NEB-241-PUS.ST25.txt

cccgtcagat	cgatatatct	ggcatcaatc	gccgaagtaa	aatttcggac	aaaatcatta	360
aatgcatcca	gccc aaataa	cgttgagata	agttcagttc	gtcttgccgg	ggccagcgcc	420
gctattcttg	agaagttgtc	aattcggttt	ttttcaacaa	agcaaaagcg	gtgctgtgct	480
tcgttatgct	caattgctaa	atcctgtcct	tgctctccta	cgccagtaat	tacaggtgca	540
gaaaactgat	cgacatgtgc	atttctaaaa	tagtcgggtt	gattacgaaa	acgcttacta	600
tcagcctcag	ctacgtacc	cagtaatgta	tattcaagcg	cttcgcagaa	actggacttc	660
ccggtaccat	tggggccata	aatcagcacc	agacgcgaat	ccaggtcaaa	ttcctcctgt	720
ctggcaaadc	ctctgaacgg	tccaacggac	aacctcctga	gtcgattgaa	agtggagacg	780
cgttcgttgc	tttggttcagg	cagtggctgg	acttcaggc	tgagggtatc	ccaggcaggt	840
tgcgccagat	cgacgatacg	ccttatccgc	tgtccctgtg	aggtacctaa	cgggataata	900
ttatccagat	tatcccatac	aagattcgcc	atTTTTctga	catcaccggg	tatatctgct	960
gtgtctaaag	tttg gaaaaa	gcgtaaaaac	tcgttactga	gcattatgaa	tcctttttta	1020
cttgctggtt	tctcacgtta	taagacaatg	ataaaagata	cactcttagc	taacgtattc	1080
acgtgatctg	tagatcaatt	atcttcagtt	ccgctctcaa	gctgaactga	accgggatga	1140
agacgggtatg	gcgcttgcca	cactagtaca	ggtgtattac	taaaaaaccg	aaaggatttc	1200
gataaagccg	attacaacgc	gttggtggac	aacaccgaag	ccacgctcgg	cgatgaactg	1260
gtggcaaaga	aagaaataca	ggtccgccgg	gagtaaactg	ccaccttcac	caagccgatg	1320
gatgagcagg	cgtaatatgt	cgcagtgtct	gcgaagcgcc	gtactccgga	tgtgcgcaag	1380
aacgactgac	gtctggtact	gagccgtgac	gatctggcct	ctgatgggcc	cgcattaatg	1440
agatggtaaa	tcctcactaa	tattgaaggc	aaaaaataaa	ggtctccaaa	atcgactctt	1500
gtaaagaggc	ttgagaggcc	ctcctgcact	ctagccatag	ttcggaattg	gtcgttaaaa	1560
tgtcgtacac	taccatcatt	ttaaaatcga	aatggaatat	tgaatggcga	acgaacgcaa	1620
aacagaatcc	ttagttcgag	accagctacg	gacatttggc	tactacgaac	cggacaacgg	1680
catttctgta	gaggagcaaa	agtccgagat	tgtcaagatt	aagggtttgc	tttcaaaagc	1740
aagtaagaac	gccaagggca	atattgggtta	tcccagattc	atcatctcta	accggaaaga	1800
tactgcattc	ctgatagtgt	tggagtgcaa	gccggatgtg	aaaaagcacg	agagcccaag	1860
ccgtgataag	ccggtagact	atgcggtgga	tggcgttctc	cactacgcca	gacacctagc	1920
caagcactat	accgtattgg	cggtggctgt	gagcggcacg	acggcaagtt	ctatgaaggt	1980
gtccaacttc	cttggtgcctg	cgggtaccac	ggatgtgaag	gcgctggtca	acgagagtaa	2040
ttcctcagtt	gccgaattgg	tgccttatga	tgactactac	cgcctggcgt	cttatgatcc	2100
ggatgttgct	cagaagcgcc	actctgactt	gctggcggtc	tcacgcgagc	tgcacgagtt	2160
tatttggacg	aaggcaaaaa	tctccgaaga	agaaaagcct	ctgctggtga	gtgggacctt	2220

NEB-241-PUS.ST25.txt

gattgcgttg	atgaacaaca	cattcatcaa	gacctttgac	gctctacctg	cagaagatgt	2280
gcaggaagcg	tggctgacgg	ctatcaagaa	ggagctggac	aaagcttcta	tccccagggc	2340
caagaaggac	acgatgctgc	agccgtatac	gacgattgcg	gttaatccca	atcttggtcaa	2400
gcctgacagc	aagacggcta	aagagtatcc	agatggagtt	ttcaaggaaa	taatcacccg	2460
catcgccgac	aacgtctggc	cctacatcaa	tgtctttcac	gactttgatg	tggtcggaca	2520
attctacggt	gagtttctga	aatatactgc	gggcgacaaa	aaagcgctgg	gcatcgtgct	2580
gacgccgcgc	catgtggctg	aactgtttctc	gctcatcgcc	aacgttaacc	ccaagtctaa	2640
ggtgctggac	atctgtgctg	gcacgggctg	ctttctcatc	tcggccatgc	aacacatgct	2700
caagaaggcc	gtaacggaca	aagagcgcaa	cgacatcaag	caaaatcggc	tcatcgggat	2760
tgaaaacaac	cccaagatgt	ttgccttggc	tgccagcaac	atgattctgc	gtggtgatgg	2820
taaggctaac	ctgcaccagg	ccagttgctt	tgataatgca	gtgattgcgg	ccgtgcagaa	2880
gatgaagccc	aacgtgggca	tgcttaaccc	cccgtattcg	cagtccaaga	gcgacgcgga	2940
actgcatgag	ctgtatttcg	tcaagcaaat	gctcgacacg	cttacaccag	gtggagttgg	3000
tatcgcgatt	gttcccatgt	caagcgccat	ctcgcccaac	ccaatgcgtg	aagagctgat	3060
gaagtaccac	tcactggatg	cggtcatgtc	aatgccccag	gagctgtttt	atccagtggg	3120
cacggtcacc	tgtgtcatgg	tctggattgc	cggtgtgcca	catgagcaaa	tgtccaagaa	3180
gacatggttt	ggctactggc	gcgacgatgg	ctttgtgaaa	accaagcata	aggggcgcat	3240
cgacatgaat	ggcacctggc	cagacatccg	tgaccgatgg	attgaaatgt	atcgcaatcg	3300
cgaagtgcac	gctggcgaga	gcatcatgca	gaaggtaggc	cccgatgatg	aatggtgcgc	3360
tgaagcctat	atggaaacgg	actactcagt	gctgactcag	tccgactttg	agaaggctcg	3420
tcaaagctac	gcgctattta	aactatttgg	tcaaggcagt	agccagtccg	aagtgaaagg	3480
ggcaacggat	gccgaagatt	aacgaccttt	ttcatctgga	gtacggtcac	agcctggagt	3540
tgaaccggct	agagcaatcc	acagcagccg	atgccgtcaa	cttcgttgga	cgggcagcta	3600
ggaacaatgg	agtcaccgca	cgctgggctc	cccctccaaa	cttgaaaccg	gcagccgcag	3660
gcaccatcag	cgtagcgctg	ggagggcaag	gtggcgcagg	agtcgccttc	ctccaaccgc	3720
gtccctactt	ttgtggccgc	gatgtgatgg	tgctgacccc	caagaagcac	atgacagacc	3780
aagaaaagct	gtggtgggtc	atgtgcatca	cagccaaccg	tttccgcttt	ggatttggtc	3840
gccaagctaa	tcggacgcta	aaggacttga	atctgcctgc	gccccaaaaa	actccaagct	3900
gggtgcatac	agcgaacccc	gatgcctacc	aagggtgtcag	gtcccccgca	agtgttcatac	3960
cagtcggcac	gctggctgtg	agcaactgga	aggctttcat	tcttcaagac	ttgtttacca	4020
tccgtaaagg	acagcgactc	accaaggcca	acatgttgcc	cggtacgggtg	ccctacatcg	4080

NEB-241-PUS.ST25.txt

gcgcatacga	cacttccaac	ggcgttactg	cgcacatcgg	gcaaaaacca	atccacgagg	4140
gcggcaccat	cagcgtcaca	tatgacgggt	caatagctga	agcgttttac	cagccctccc	4200
cattttgggc	atcggatgct	gtgaacgtgc	tctatcccaa	gggtttcaca	ctcacaccgg	4260
ccactgcctt	gtttatctgc	gcaatcatca	ggatggagaa	atatcgcttc	aactatggcc	4320
gaaaatggca	cttagagcgt	atgcgagaga	cagttatcag	gttaccagct	actgcaacag	4380
gtgcaccaga	ttgggacttt	atggagaaat	acatcaaaac	tttgccctat	agctcgagct	4440
tgcaataatc	atggctgatt	tcctaaattt	cctgccgcat	ctacgggtat	tgcatgttca	4500
ggacggtggt	gatcatcgct	aggtggaggc	ggaaagccgt	gttttgctga	ccgcttgccc	4560
ggcctgcggt	gaaaagcctt	cccattcagg	gaaggcttta	atcgagttat	agatct	4616

<210> 6
 <211> 1899
 <212> DNA
 <213> unknown

<220>
 <223> restriction and modification system of Citrobacter species 2144

<400> 6	
atggcgaacg	60
aacgcaaaac	
agaatcctta	
gttcgagacc	
agctacggac	
atttggttac	
tacgaaccgg	120
acaacggcat	
ttctgtagag	
gagcaaaagt	
ccgagattgt	
caagattaag	
ggtttgcttt	180
caaaagcaag	
taagaacgcc	
aagggcaata	
ttggttatcc	
cgagttcatc	
atctctaacc	240
ggaaagatac	
tgcatctctg	
atagttgtgg	
agtgaagcc	
ggatgtgaaa	
aagcacgaga	300
gccaagccg	
tgataagccg	
gtagactatg	
cggtggatgg	
cgttctccac	
tacgccagac	360
acctagccaa	
gcactatacc	
gtattggcgg	
tggtgtgag	
cggcacgacg	
gcaagttcta	420
tgaaggtgtc	
caacttcctt	
gtgcctgcgg	
gtaccacgga	
tgtgaaggcg	
ctggtcaacg	480
agagtaattc	
ctcagttgcc	
gaattggtgc	
cttatgatga	
ctactaccgc	
ctggcgtctt	540
atgatccgga	
tgttgctcag	
aagcgccact	
ctgacttgct	
ggcgttctca	
cgcgagctgc	600
acgagtttat	
ttggacgaag	
gcaaaaatct	
ccgaagaaga	
aaagcctctg	
ctggtgagtg	660
ggaccttgat	
tgcgttgatg	
aacaacacat	
tcatcaagac	
ctttgacgct	
ctacctgcag	720
aagatgtgca	
ggaagcgtgg	
ctgacggcta	
tcaagaagga	
gctggacaaa	
gcttctatcc	780
cccaggccaa	
gaaggacacg	
atgctgcagc	
cgtatacgac	
gattgcgggt	
aatcccaatc	840
ttggcaagcc	
tgacagcaag	
acggctaaag	
agtatccaga	
tgaggttttc	
aaggaaataa	900
tcacccgcat	
cgccgacaac	
gtctggccct	
acatcaatgt	
ctttcacgac	
tttgatgtgg	960
tcggacaatt	
ctacgggtgag	
tttctgaaat	
atactgcggg	
cgacaaaaaa	
gcgctgggca	1020
tcgtgctgac	
gccgcgccat	
gtggctgaac	
tgttctcgct	
catcgccaac	
gttaacccca	1080
agtctaaggt	
gctggacatc	
tgtgcgggca	
cgggaggctt	
tctcatctcg	

NEB-241-PUS.ST25.txt

gccatgcaac	acatgctcaa	gaaggccgta	acggacaaag	agcgcaacga	catcaagcaa	1140
aatcggtca	tctggattga	aaacaacccc	aagatgtttg	ccttggctgc	cagcaacatg	1200
attctgcgtg	gtgatggtaa	ggctaacctg	caccaggcca	gttgctttga	taatgcagtg	1260
attgcgccg	tgcagaagat	gaagcccaac	gtgggcatgc	ttaaccccc	gtattcgag	1320
tccaagagcg	acgcggaact	gcatgagctg	tatttcgtca	agcaaagtct	cgacacgctt	1380
acaccaggtg	gagttggtat	cgcgattggt	cccatgtcaa	gcgccatctc	gccaaccca	1440
atgctgaag	agctgatgaa	gtaccactca	ctggatgcgg	tcatgtcaat	gccccaggag	1500
ctgttttatc	cagtgggcac	ggtcacctgt	gtcatggctt	ggattgccgg	tgtgccacat	1560
gagcaaatgt	ccaagaagac	atggtttggc	tactggcgcg	acgatggctt	tgtgaaaacc	1620
aagcataagg	ggcgcatcga	catgaatggc	acctggccag	acatccgtga	ccgatggatt	1680
gaaatgtatc	gcaatcgcg	agtgcagtct	ggcgagagca	tcatgcagaa	ggtaggcccc	1740
gatgatgaat	ggtgcgctga	agcctatatg	gaaacggact	actcagtgtc	gactcagtcc	1800
gactttgaga	aggtcgttca	aagctacgcg	ctatttaaac	tatttgggtca	aggcagtagc	1860
cagtccgaag	tgaaaggggc	aacggatgcc	gaagattaa			1899

<210> 7
 <211> 960
 <212> DNA
 <213> unknown

<220>
 <223> specificity subunit of Citrobacter species 2144

<400> 7						
atgccgaaga	ttaacgacct	ttttcatctg	gagtacggtc	acagcctgga	gttgaaccgg	60
ctagagcaat	ccacagcagc	cgatgccgtc	aacttcgttg	gacgggcagc	taggaacaat	120
ggagtcaccg	cacgcgtggc	tccccctcca	aacttgaaac	cggcagccgc	aggcaccatc	180
agcgtagcgc	tgggagggca	aggtggcgca	ggagtcgcct	tcctccaacc	gcgtccctac	240
ttttgtggcc	gcgatgtgat	ggtgctgacc	ccaagaagc	acatgacaga	ccaagaaaag	300
ctgtggtggg	tcatgtgcat	cacagccaac	cgtttccgct	ttggatttgg	tcgccaagct	360
aatcggacgc	taaaggactt	gaatctgcct	gcgccccaaa	aaactccaag	ctgggtgcat	420
acagcgaacc	ccgatgccta	ccaagggtgc	aggtcccccg	caagtgttca	tccagtcggc	480
acgctggctg	tgagcaactg	gaaggctttc	attcttcaag	acttgtttac	catccgtaaa	540
ggacagcgac	tcaccaaggc	caacatgttg	cccggtagcg	tgccctacat	cggcgcatcg	600
gacacttcca	acggcgttac	tgcgcacatc	gggcaaaaac	caatccacga	gggcggcacc	660
atcagcgtca	catatgacgg	ttcaatagct	gaagcgtttt	accagccctc	cccattttgg	720

NEB-241-PUS.ST25.txt

gcacgcatg ctgtgaacgt gctctatccc aagggtttca cactcacacc ggccactgcc 780
 ttgtttatct gcgcaatcat caggatggag aaatatcgct tcaactatgg ccgaaaatgg 840
 cacttagagc gtatgcgaga gacagttatc aggttaccag ctactgcaac aggtgcacca 900
 gattgggact ttatggagaa atacatcaaa actttgcct atagctcgca gttgcaataa 960

<210> 8
 <211> 632
 <212> PRT
 <213> unknown

<220>
 <223> predicted amino acid sequence of restriction modification system
 of Citrobacter species 2144

<400> 8

Met Ala Asn Glu Arg Lys Thr Glu Ser Leu Val Arg Asp Gln Leu Arg
 1 5 10 15

Thr Phe Gly Tyr Tyr Glu Pro Asp Asn Gly Ile Ser Val Glu Glu Gln
 20 25 30

Lys Ser Glu Ile Val Lys Ile Lys Gly Leu Leu Ser Lys Ala Ser Lys
 35 40 45

Asn Ala Lys Gly Asn Ile Gly Tyr Pro Glu Phe Ile Ile Ser Asn Arg
 50 55 60

Lys Asp Thr Ala Phe Leu Ile Val Val Glu Cys Lys Pro Asp Val Lys
 65 70 75 80

Lys His Glu Ser Pro Ser Arg Asp Lys Pro Val Asp Tyr Ala Val Asp
 85 90 95

Gly Val Leu His Tyr Ala Arg His Leu Ala Lys His Tyr Thr Val Leu
 100 105 110

Ala Val Ala Val Ser Gly Thr Thr Ala Ser Ser Met Lys Val Ser Asn
 115 120 125

Phe Leu Val Pro Ala Gly Thr Thr Asp Val Lys Ala Leu Val Asn Glu
 130 135 140

Ser Asn Ser Ser Val Ala Glu Leu Val Pro Tyr Asp Asp Tyr Tyr Arg
 145 150 155 160

Leu Ala Ser Tyr Asp Pro Asp Val Ala Gln Lys Arg His Ser Asp Leu
 165 170 175

NEB-241-PUS.ST25.txt

Leu Ala Phe Ser Arg Glu Leu His Glu Phe Ile Trp Thr Lys Ala Lys
 180 185 190
 Ile Ser Glu Glu Glu Lys Pro Leu Leu Val Ser Gly Thr Leu Ile Ala
 195 200 205
 Leu Met Asn Asn Thr Phe Ile Lys Thr Phe Asp Ala Leu Pro Ala Glu
 210 215 220
 Asp Val Gln Glu Ala Trp Leu Thr Ala Ile Lys Lys Glu Leu Asp Lys
 225 230 235 240
 Ala Ser Ile Pro Gln Ala Lys Lys Asp Thr Met Leu Gln Pro Tyr Thr
 245 250 255
 Thr Ile Ala Val Asn Pro Asn Leu Gly Lys Pro Asp Ser Lys Thr Ala
 260 265 270
 Lys Glu Tyr Pro Asp Gly Val Phe Lys Glu Ile Ile Thr Arg Ile Ala
 275 280 285
 Asp Asn Val Trp Pro Tyr Ile Asn Val Phe His Asp Phe Asp Val Val
 290 295 300
 Gly Gln Phe Tyr Gly Glu Phe Leu Lys Tyr Thr Ala Gly Asp Lys Lys
 305 310 315 320
 Ala Leu Gly Ile Val Leu Thr Pro Arg His Val Ala Glu Leu Phe Ser
 325 330 335
 Leu Ile Ala Asn Val Asn Pro Lys Ser Lys Val Leu Asp Ile Cys Ala
 340 345 350
 Gly Thr Gly Gly Phe Leu Ile Ser Ala Met Gln His Met Leu Lys Lys
 355 360 365
 Ala Val Thr Asp Lys Glu Arg Asn Asp Ile Lys Gln Asn Arg Leu Ile
 370 375 380
 Gly Ile Glu Asn Asn Pro Lys Met Phe Ala Leu Ala Ala Ser Asn Met
 385 390 395 400
 Ile Leu Arg Gly Asp Gly Lys Ala Asn Leu His Gln Ala Ser Cys Phe
 405 410 415
 Asp Asn Ala Val Ile Ala Ala Val Gln Lys Met Lys Pro Asn Val Gly
 420 425 430

NEB-241-PUS.ST25.txt

Met Leu Asn Pro Pro Tyr Ser Gln Ser Lys Ser Asp Ala Glu Leu His
435 440 445

Glu Leu Tyr Phe Val Lys Gln Met Leu Asp Thr Leu Thr Pro Gly Gly
450 455 460

Val Gly Ile Ala Ile Val Pro Met Ser Ser Ala Ile Ser Pro Asn Pro
465 470 475 480

Met Arg Glu Glu Leu Met Lys Tyr His Ser Leu Asp Ala Val Met Ser
485 490 495

Met Pro Gln Glu Leu Phe Tyr Pro Val Gly Thr Val Thr Cys Val Met
500 505 510

Val Trp Ile Ala Gly Val Pro His Glu Gln Met Ser Lys Lys Thr Trp
515 520 525

Phe Gly Tyr Trp Arg Asp Asp Gly Phe Val Lys Thr Lys His Lys Gly
530 535 540

Arg Ile Asp Met Asn Gly Thr Trp Pro Asp Ile Arg Asp Arg Trp Ile
545 550 555 560

Glu Met Tyr Arg Asn Arg Glu Val His Ala Gly Glu Ser Ile Met Gln
565 570 575

Lys Val Gly Pro Asp Asp Glu Trp Cys Ala Glu Ala Tyr Met Glu Thr
580 585 590

Asp Tyr Ser Val Leu Thr Gln Ser Asp Phe Glu Lys Val Val Gln Ser
595 600 605

Tyr Ala Leu Phe Lys Leu Phe Gly Gln Gly Ser Ser Gln Ser Glu Val
610 615 620

Lys Gly Ala Thr Asp Ala Glu Asp
625 630

<210> 9
<211> 319
<212> PRT
<213> unknown

<220>
<223> predicted amino acid sequence of the specificity subunit of
Citrobacter species 2144

<400> 9

Met Pro Lys Ile Asn Asp Leu Phe His Leu Glu Tyr Gly His Ser Leu
1 5 10 15

Glu Leu Asn Arg Leu Glu Gln Ser Thr Ala Ala Asp Ala Val Asn Phe
20 25 30

Val Gly Arg Ala Ala Arg Asn Asn Gly Val Thr Ala Arg Val Ala Pro
35 40 45

Pro Pro Asn Leu Lys Pro Ala Ala Ala Gly Thr Ile Ser Val Ala Leu
50 55 60

Gly Gly Gln Gly Gly Ala Gly Val Ala Phe Leu Gln Pro Arg Pro Tyr
65 70 75 80

Phe Cys Gly Arg Asp Val Met Val Leu Thr Pro Lys Lys His Met Thr
85 90 95

Asp Gln Glu Lys Leu Trp Trp Val Met Cys Ile Thr Ala Asn Arg Phe
100 105 110

Arg Phe Gly Phe Gly Arg Gln Ala Asn Arg Thr Leu Lys Asp Leu Asn
115 120 125

Leu Pro Ala Pro Gln Lys Thr Pro Ser Trp Val His Thr Ala Asn Pro
130 135 140

Asp Ala Tyr Gln Gly Val Arg Ser Pro Ala Ser Val His Pro Val Gly
145 150 155 160

Thr Leu Ala Val Ser Asn Trp Lys Ala Phe Ile Leu Gln Asp Leu Phe
165 170 175

Thr Ile Arg Lys Gly Gln Arg Leu Thr Lys Ala Asn Met Leu Pro Gly
180 185 190

Thr Val Pro Tyr Ile Gly Ala Ser Asp Thr Ser Asn Gly Val Thr Ala
195 200 205

His Ile Gly Gln Lys Pro Ile His Glu Gly Gly Thr Ile Ser Val Thr
210 215 220

Tyr Asp Gly Ser Ile Ala Glu Ala Phe Tyr Gln Pro Ser Pro Phe Trp
225 230 235 240

Ala Ser Asp Ala Val Asn Val Leu Tyr Pro Lys Gly Phe Thr Leu Thr
Page 10

245

255

Pro Ala Thr Ala Leu Phe Ile Cys Ala Ile Ile Arg Met Glu Lys Tyr
260 265 270

Arg Phe Asn Tyr Gly Arg Lys Trp His Leu Glu Arg Met Arg Glu Thr
275 280 285

Val Ile Arg Leu Pro Ala Thr Ala Thr Gly Ala Pro Asp Trp Asp Phe
290 295 300

Met Glu Lys Tyr Ile Lys Thr Leu Pro Tyr Ser Ser Gln Leu Gln
305 310 315

<210> 10
<211> 23
<212> DNA
<213> unknown

<220>
<223> synthetic

<400> 10
cagagagata acccacaaga ttg

23

<210> 11
<211> 24
<212> DNA
<213> unknown

<220>
<223> synthetic

<400> 11
ccacaagaat tgagttaagc ccaa

24

<210> 12
<211> 25
<212> DNA
<213> unknown

<220>
<223> synthetic

<400> 12
atcgagagat aaccacaag aattg

25

<210> 13
<211> 25
<212> DNA
<213> unknown

<220>
<223> synthetic

<220>
 <221> misc_feature
 <222> (4)..(8)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (13)..(25)
 <223> n=a,c,t or g

<400> 13
 caannnnngt ggnnnnnnnn nnnnn

25

<210> 14
 <211> 12
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (4)..(8)
 <223> n=a,t,c or g

<400> 14
 caannnnngt gg

12

<210> 15
 <211> 21
 <212> DNA
 <213> unknown

<220>
 <223> primer

<400> 15
 cagttcgatg taaccactc g

21

<210> 16
 <211> 23
 <212> DNA
 <213> unknown

<220>
 <223> primer

<400> 16
 cccgctgacg cgccctgacg ggc

23

<210> 17
 <211> 43
 <212> DNA
 <213> unknown

<220>

<223> synthetic

<400> 17

aagtgccacc tgacgtgcaa cctaggtggc acgtctaaga aac

43

<210> 18

<211> 43

<212> DNA

<213> unknown

<220>

<223> synthetic

<400> 18

gtttcttaga cgtgccacct aggttgacg tcaggtggca ctt

43

<210> 19

<211> 44

<212> DNA

<213> unknown

<220>

<223> synthetic

<400> 19

tggtttctta gacgtgccac ctaggttgca cgtcaggtgg cact

44

<210> 20

<211> 42

<212> DNA

<213> unknown

<220>

<223> synthetic

<400> 20

tgccacctga cgtgcaacct aggtggcacg tctaagaaac ca

42

<210> 21

<211> 43

<212> DNA

<213> unknown

<220>

<223> synthetic

<400> 21

agtgccacct gacgtgcaac ctaggtggca cgtctaagaa acc

43

<210> 22

<211> 43

<212> DNA

<213> unknown

<220>

<223> synthetic

<400> 22

agtgccacct gacgtgccac ccgggttgca cgtctaagaa acc

43

<210> 23
 <211> 18
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> n=a,c,t or g

<400> 23
 gcaaccnggg tggcacgt

18

<210> 24
 <211> 11
 <212> PRT
 <213> unknown

<220>
 <223> synthetic

<400> 24

Val Leu Asp Ile Cys Ala Gly Thr Gly Gly Phe
 1 5 10

<210> 25
 <211> 10
 <212> PRT
 <213> unknown

<220>
 <223> synthetic

<400> 25

Ala Asn Glu Arg Lys Thr Glu Glu Leu Val
 1 5 10

<210> 26
 <211> 11
 <212> PRT
 <213> unknown

<220>
 <223> synthetic

<400> 26

Met Ala Asn Glu Arg Lys Thr Glu Ser Leu Val
 1 5 10

<210> 27
 <211> 10
 <212> PRT
 <213> unknown

<220>
 <223> synthetic

<400> 27

Pro Lys Ile Asn Asp Leu Phe His Leu Glu
 1 5 10

<210> 28
 <211> 11
 <212> PRT
 <213> unknown

<220>
 <223> synthetic

<400> 28

Met Pro Lys Ile Asn Asp Leu Phe His Leu Glu
 1 5 10

<210> 29
 <211> 12
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (5)..(8)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> n is a, c, g, or t

<400> 29
 ccacnnnnnt tg

12

<210> 30
 <211> 36
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (1)..(11)

<223> n=a,c,t or g

<220>

<221> misc_feature

<222> (15)..(19)

<223> n=a,c,t or g

<220>

<221> misc_feature

<222> (24)..(36)

<223> n=a,c,t or g

<400> 30

nnnnnnnnnn ncaannnnng tggnnnnnnn nnnnnn

36

<210> 31

<211> 34

<212> DNA

<213> unknown

<220>

<223> synthetic

<220>

<221> misc_feature

<222> (1)..(10)

<223> n=a,c,g or t

<220>

<221> misc_feature

<222> (14)..(18)

<223> n=a,c,g or t

<220>

<221> misc_feature

<222> (23)..(34)

<223> n=a,c,g or t

<400> 31

nnnnnnnnnn caannnnngt ggnnnnnnnn nnnn

34

<210> 32

<211> 35

<212> DNA

<213> unknown

<220>

<223> synthetic

<220>

<221> misc_feature

<222> (1)..(10)

<223> n=a,c,t or g

<220>

<221> misc_feature

<222> (14)..(18)

<223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (23)..(35)
 <223> n=a,c,t or g

<400> 32
 nnnnnnnnnn caannnnngt ggnnnnnnnn nnnnn

35

<210> 33
 <211> 35
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (1)..(11)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (15)..(19)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (24)..(35)
 <223> n=a,c,t or g

<400> 33
 nnnnnnnnnn ncaannnnng tggnnnnnnn nnnnn

35

<210> 34
 <211> 13
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (4)..(9)
 <223> n=a,c,t or g

<400> 34
 caannnnngt tgg

13

<210> 35
 <211> 11
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
<221> misc_feature
<222> (4)..(7)
<223> n=a,c,g or t

<400> 35
caannnngtg g

11

<210> 36
<211> 11
<212> DNA
<213> unknown

<220>
<223> synthetic

<220>
<221> misc_feature
<222> (4)..(8)
<223> n=a,c,t or g

<400> 36
caannnnntt g

11

<210> 37
<211> 13
<212> DNA
<213> unknown

<220>
<223> synthetic

<220>
<221> misc_feature
<222> (5)..(9)
<223> n=a,c,t or g

<400> 37
ccacnnnnng tgg

13

<210> 38
<211> 12
<212> DNA
<213> unknown

<220>
<223> synthetic

<220>
<221> misc_feature
<222> (4)..(4)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (5)..(8)

<223> n=a,c,t or g

<220>

<221> misc_feature

<222> (12)..(12)

<223> r=a or g

<400> 38

caannnnngt gr

12

<210> 39

<211> 11

<212> DNA

<213> unknown

<220>

<223> synthetic

<220>

<221> misc_feature

<222> (4)..(8)

<223> n=a,c,t or g

<400> 39

caannnnngt g

11

<210> 40

<211> 11

<212> DNA

<213> unknown

<220>

<223> synthetic

<220>

<221> misc_feature

<222> (4)..(8)

<223> n=a,c,t or g

<400> 40

cgannnnntg c

11

<210> 41

<211> 11

<212> DNA

<213> unknown

<220>

<223> synthetic

<220>

<221> misc_feature

<222> (4)..(8)

<223> n=a,c,t or g

<400> 41

caannnnntg c

11

<210> 42
 <211> 12
 <212> DNA
 <213> unknown

<220>
 <223> synthetic

<220>
 <221> misc_feature
 <222> (4)..(8)
 <223> n=a,c,t or g

<400> 42
 gcannnnngt gg 12

<210> 43
 <211> 12
 <212> DNA
 <213> Bacillus coagulans

<220>
 <221> misc_feature
 <222> (4)..(9)
 <223> n=a,c,t or g

<400> 43
 cgannnnnnt gc 12

<210> 44
 <211> 11
 <212> DNA
 <213> Bacillus sphaericus

<220>
 <221> misc_feature
 <222> (3)..(6)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (10)..(10)
 <223> y=c or t

<400> 44
 acnnngtay c 11

<210> 45
 <211> 11
 <212> DNA
 <213> Bacillus pumilus

<220>
 <221> misc_feature

<222> (4)..(8)
 <223> n=a,c,g or t

<400> 45
 gagnnnnnct c

11

<210> 46
 <211> 11
 <212> DNA
 <213> Campylobacter jejuni

<220>
 <221> misc_feature
 <222> (4)..(9)
 <223> n=a, c, t or g

<400> 46
 ccannnnnng t

11

<210> 47
 <211> 13
 <212> DNA
 <213> Acinetobacter lwoffii

<220>
 <221> misc_feature
 <222> (5)..(10)
 <223> n=a, c, t or g

<400> 47
 gaacnnnnnn tcc

13

<210> 48
 <211> 11
 <212> DNA
 <213> Haemophilus aegyptius

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> y=c or t

<220>
 <221> misc_feature
 <222> (4)..(8)
 <223> n=a,c,t or g

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> r=a or g

<400> 48
 gaynnnnnrt c

11

<210> 49

NEB-241-PUS.ST25.txt

<211> 11
<212> DNA
<213> Bacillus stearothermophilus

<220>
<221> misc_feature
<222> (3)..(7)
<223> n=a,c, t or g

<400> 49
acnnnnnctc c

11